AD-A228 949

## REPORT DOCUMENTATION PAGE

DTIC FILE COPY

1b. RESTRICTIVE MARKINGS: NONE

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE: N/A

3. DISTRIBUTION / AVAILABILITY OF REPORT: unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S): 0388-1

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 90 1097

6a. NAME OF PERFORMING ORGANIZATION: Univ. of Maryland

6b. OFFICE SYMBOL (If applicable):

7a. NAME OF MONITORING ORGANIZATION: AFOSR

6c. ADDRESS (City, State, and ZIP Code): College Park, MD 20742

7b. ADDRESS (City, State, and ZIP Code): Bolling AFB, Washington, DC 20332-5260

8a. NAME OF FUNDING / SPONSORING ORGANIZATION: AFOSR

8b. OFFICE SYMBOL (If applicable): NE

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0388

8c. ADDRESS (City, State, and ZIP Code): Bolling AFB, Washington, DC 20332-5260

10. SOURCE OF FUNDING NUMBERS

| PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO. |
|---|---|---|---|
| 61102F | 2305 | B1 | |

11. TITLE (Include Security Classification): Connectionist Models for Intelligent Computation

12. PERSONAL AUTHOR(S): H.H. Chen and Y.C. Lee

13a. TYPE OF REPORT: Annual Technical

13b. TIME COVERED: FROM 9/1/88 TO 8/31/89

14. DATE OF REPORT (Year, Month, Day): 7/26/89

15. PAGE COUNT:

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

| FIELD | GROUP | SUB-GROUP |
|---|---|---|
| | | |
| | | |

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)


In the past year, we have 1) continued our investigation in the high-order neural network architectures 2) constructed a stereopsis network that is learned analytically and 3) proposed a novel scheme (PSIN) to automatically build up the network while learning.




20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION: Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL: Giles

22b. TELEPHONE (Include Area Code): 202-767-4931

22c. OFFICE SYMBOL: NE

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE: Unclassified

Project Title: Connectionist Models for Intelligent Computation

Contract/Grant No.: AFOSR-87-0388

Contract/Grant Period of Performance: Sept. 1, 1988-Aug. 31, 1989

Principal Investigator: H. H. Chen Telephone: (301) 454-3182
Y. C. Lee (301) 454-3177

Research Objective:

To study the underlying principles, architectures and appilications of artificial neural networks for intelligent computations.

Approach:

We use both numerical simulation and theoretical analysis to investigate various alternatives in connection schemes, organization principles and architectures of aritificial neural networks.

Progress for period 9/1/88-8/31/89:

In the past year, our research on neural network models for intelligent computing under the sponsorship of AFOSR continued to make important progress. In particular, we have constructed the Parallel Sequential Induction Network, a powerful network that self-organizes into an optimal structure to perform classification tasks.

In neural network research, much attention has been paid to improving the efficiency of learning connection weights for a network with fixed topology. However, little progress has been made toward uncovering optimal designing principles to reshape the connection topology of a network adaptively to maximize the performance of a specific task. Recent studies indicate that multi-layered feedforward networks of sufficient complexity, in general, need only two hidden layers to imitate any decision hypersurface in the pattern space. However, little is known about the learning process that takes place in the formation of these decision regions. In fact, we can well imagine the difficulty the network has to face in forming all these widely scattered disjoint decision regions necessary for the requisite task. Only input patterns that reside in the nearby locations could have a strong positive influence on the learning of their formation. The overwhelming majority of input patterns that lies outside this region would only contribute large noises and may cause failure for the network to learn properly.

We conceived a solution to this problem by combining the best of both the parallel and the sequential strategies to optimize the performance of a neural network classifier. First, we took

(KR)

a parallel approach by assigning an output decision neuron to each decision region in the pattern space. Instead of relying on a single decision neuron to carry the full burden of making the complex decision by itself, all the output decision neurons at each decision region would participate to share the arduous burden of figuring out the much simpler decision boundaries at their own respective local neighborbood. For these local decision neurons, there is no need to use a complicated network; often, simple perceptions would suffice for the job. These local decision neurons are, in a way, very similar to hidden neurons in a feedforward network. The main difference between them is that the former are not hidden and they can be directly trained without the need to backpropagate the errors through the hidden layers—a task which often causes the learning to slow down drastically.

The process of training the neural network is to reduce the intercategorical mixing of these input patterns received by each decision neuron. We proposed to use the mutual information entropy as the objective function to be maximized to reduce optimally the mixing of the input patterns. A stochastic gradient descent algorithm is derived for the optimization task. This self-organizing training with top-down inputs is found to be very efficient in solving complex decision problems.

Even though the above parallel self-orginization scheme for the connection topology could be made very efficient in solving difficult classification problems with complex decision boundaries by breaking it down into many small regions having simple boundaries, it could be (due to an insufficient number of decision neurons allocated or other reasons) that the learned networks are still not optimal in their performance. Our second strategy is to use a sequence of simple networks added on top of the existing network structures. The use of mutual information entropy as the objective function and the inherent self-organization behavior makes this procedure a natural choice. In general, a decision tree of networks will be grown to optimize the performance. However, since each network node in the tree is already an optimized parallel network, the trees thus obtained are usually small.

Although artificial neural networks have enjoyed a great deal of success lately in a limited domain, the inherent deficiencies of current network models may ultimately prove too restrictive for the kinds of applications these models are originally intended for. One of the major shortcomings of the current neural networks is the inability of the networks to model the higher level human reasoning processes which are generally thought to require the expressiveness of a Turing machine. Most of the current generation of neural network architectures can be identified as the feedforward or nonrecursive kind. As such, they are essentially a kind of adaptive function estimator.

Hence, their expressive power is necessarily extremely limited. In order to enhance their expressive power, it is imperative to add recursiveness and a working memory to the structure. In the coming years, we intend to pursue the study of learning procedures, the search of pertinent recurrent connections, and the interconnection between the neural network finite machine and the working memory.

Publications

1. Lee, Y. C., "Efficient Stochastic Gradient Descent Learning Algorithm for Neural Network." University of Maryland Preprint: CS-TR-1973, 1988.

2. Maxwell T., C. L. Giles and Y. C. Lee, "Transformation Invariance with High Order Correlation in Neural Net Architecture." University of Maryland Preprint: UMLPF 87-027.

3. Sun, G. Z., H. H. Chen and Y. C. Lee, "Parallel Sequential Inductive Network Architecture." Proceeding, Vol. 1: 2nd International Conference on Neural Networks, San Diego, 1988, p489.

4. _________________, "Learning Decision Trees Using Parallel Sequential Induction Networks." Neural Network, Vol. 1, Supplement 1, 1988. p54.

5. _________________, "A Neural Network Approach to Speech Recognition." Neural Network, Vol. 1, Supplement 1, 1988, p306.

6. _________________, "Pattern Classifier Net Using Information Entropy." Submitted to the International Joint Conference on Neural Networks, Washington, DC, June 1989.

7. _________________, "Hand Written Letter Recognition with Neural Networks." Submitted to the International Joint Conference on Neural Networks, Washington, DC, June 1989.